Open Access
Phytopathogenomics and Disease Control., 2022, 1(2):85-89
ISSN (Online):2957-5842; ISSN (Print):2957-5834
DOI: https://doi.org/10.22194/pdc/2.17
https://societyfia.org/journal/PDC

DNA barcoding of an endangered plant species, *salvadora oleoides*

Sadia Ali* and Maria Babar

Centre of Agricultural Biochemistry and Biotechnology, University of Agriculture, Faisalabad, Pakistan
*Corresponding author's e-mail: biotech327@gmail.com

Salvadora oleoides Decne is an endangered plant species. This species is important as the plant is used for multiple purposes. Natives use the plants for their medicinal needs in treating rheumatic pains and coughs, and many others. This species is threatened by over-exploitation and deforestation or illegal logging, as well as its very slow growth rate, so the accurate identification of this species is important. Knowledge about threats or even distribution is limited, making identifying species more difficult. Therefore, this research aimed at DNA barcoding of endangered species *Salvadora oleoides* to assess their potential at the meaningful taxonomic level. This study aims to amplify the *rbc*L and *mat*K genes in *Salvadora oleoides* collected from the different regions of Pakistan. The plant DNA isolation was carried out by the CTAB method. PCR analysis was performed, which showed amplification of both *rbc*L and *mat*K regions. The amplified products were purified and sent for sequencing. The results revealed a 100% sequencing rate for *rbc*L; however, the *mat*K amplified product was not sequenced due to the high heterozygosity in this region, which might result in polymerase slippage during the Sanger sequencing method. The *rbc*L sequence was checked for homology by BLASTn and then submitted to NCBI for GenBank accession. The accession number assigned to *Salvadora oleoides* was OP046316 for *rbc*L.
Keywords: DNA Barcoding, Medicinal Plants, *matk, rbcl,* BLAST, Species Identification.

INTRODUCTION

An endangered species *Salvadora oleoides* from the Salvadoraceae family containing three genera has 12 species. This plant is used for multiple purposes, such as the oil extracted from the tree for medicinal purposes. It is distributed in Africa and subtropical and tropical areas of Asia mainly. The species is extremely valuable commercially; however, it is currently endangered due to overexploitation for wood and medicinal purposes. This plant has numerous alkaloids, glycosides, terpenoids, and flavonoids exploited for pharmacological activities (Khan *et al*., 2019).

The tree species are thought to be quickly diminishing due to a lack of suitable habitats, overexploitation, human activities, and agricultural land development (Khan, 1994). *Salvadora oleoides*, with a population size of 11, is listed as Critically Endangered and at risk of going extinct, according to a survey done in 2012 (Khan *et al*.,2012). Based on their geographic range and the number of locales found, tree species were assessed for their state of conservation. Due to its limited geographic distribution and single locality, S. oleoides was designated as being in the Endangered Category. Tree species' conservation status depends on its geographic range and the number of localities/sub-populations (Khan *et al*., 2012).

Plants of medicinal importance, such as *S. oleoides*, are typically sold in a dried form of leaves, dried roots, and tree barks at markets and traditional herbal shops or in processed powdered form in portions, mixes, or extracted material. Proper identification of the plants requires many morphological characteristics that are mostly missing for such plant material to identify by retailers and customers. Plant's Aerial parts might lose essential diagnostic features for taxonomical identification, while identification by root is usually difficult due to a lack of distinguishing morphology (Chen. *et al*., 2016). Furthermore, taxonomic identification utilizing macro- and micromorphological, as well as organoleptic approaches, can be time-taking, requires skill, and should be error-prone with accurate references (Li *et al*., 2011). Structural resemblances among some of the plant species and dried form of plant parts, the shortage of medicinally important plant species in wildlife, inconsiderate collection methods and practices, and the absence of a standard system of correct identification and control are major issues that contribute to both unintentional and deliberate replacement of species (Ghorbani *et al*., 2017).

Dara, M.Z.N., M.K. Sarwar and A. Nisar. 2022. Molecular probing of wilt etiology in *Thuja orientalis*, and pathogen confirmation through Koch's Postulates. Phytopathogenomics and Disease Control 1:85-89.
[Received 4 Jul 2022; Accepted 15Nov 2022; Published 27 Dec 2022]

DNA barcoding proved to be an efficient tool for proper discrimination of the related species by generating standers that can be used universally. The discrimination is done by accurately sequencing the standard gene region in a very short time (Hebert *et al.*, 2003).

RbcL and matK have been having most recommended and vasty studied DNA barcodes for plants, as they are proved by PWG (Pair wise group) to be potential standardized barcodes for plant DNA barcoding. These barcodes are used to properly identify and discriminate any native or foreign plant species *(*Maloukh *et al*., 2017*)*. In some cases, *rbc*L showed less discriminatory power than matK but had other advantages of universal, clear alignment, unambiguity, and higher frequency for sequencing (Dong *et al*., 2014).

MATERIALS AND METHODS

Plant material: The plant species *Salvadora oleoides* was used in the study (Fig. 1). The germplasm of *Salvadora oleoides*, was obtained from different regions of Pakistan, including District of Mianwali, Bhakker Layyeh, MuzaffarGarh, some parts of the district of Khusab, Sargodha and Jhang. The plant material used during the experiment was a leaves sample *of Salvadora oleoides.*

Figure 1. Collected sample of *Salvadora oleoides.*

DNA extraction: DNA from young leaves of *Salvadora oleoides* was extracted using the modified CTAB method (Fig. 2). Young leaf samples were collected from the plant and crushed using a sterile mortar and pestle. 1-2 ml of preheated CTAB Buffer was added and samples were grounded. To make the extracted DNA precipitate, chilled propanol was added. One µl β-Marcapto ethanol was added as the phenolic content was higher. The calculated volume of PCI and then phenol was used. After washing with 70% ethanol, The quality of extracted DNA was further estimated using Nanodrop 8000 (Spectrophotometer, Thermo SCIENTIFIC). The isolated DNA was stored at -20°C.

Figure 2. Isolated DNA of *Salvadora oleoides* .

PCR analysis: The following ingredients were used in a 50 µL PCR reaction mixture: 3 µL of template DNA (300 ng), 5 µL of PCR buffer (10X), 5 µL of $MgCl_2$ (25 mM), 5µL of forward and reverse primers (10 pmol), 0.35 µL of Taq polymerase, 5 µL of dNTPs (1 mM), and the remaining was deionized distilled water. For PCR analysis, the heat profile and reagent profile were both standardized. In Salvadora oleoides, an endangered plant species in salty environments, the chloroplasts region rbcL and matK were employed as a DNA barcode for species identification.

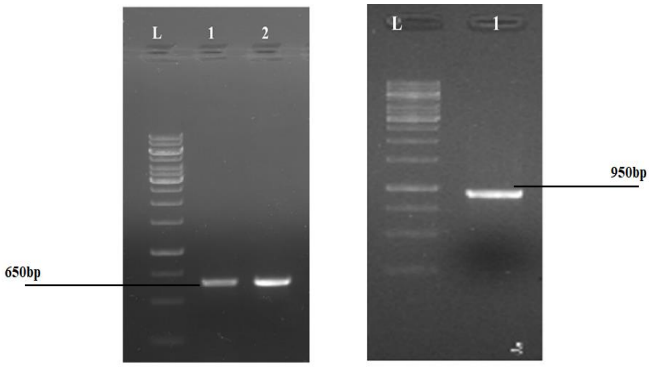
Nucleotide sequencing: The 1% TAE gel was used to extract the PCR-amplified products, which were then purified using a DNA purification kit (Thermo Scientific). The cleaned samples were again measured using a UV-2800 spectrophotometer (BMS).

The following ingredients were used to construct the 20 µL sequencing reaction: 2 µL of template DNA (20 ng), 1 µL of sequencing buffer (5X), and 2 µL of large dye terminator. The thermal profile was modified to 35 cycles, with the following temperatures: 95°C for 5 minutes, 94°C for 1 minute, 53°C for 45 seconds, 68°C for 4 minutes, and the final extension at 68°C for 10 minutes. With sterile distilled water, the final volume was changed to 20 µL.

Two volumes of 80% propanol were used to precipitate the ordered products, and they were then given two washes with 80% ethanol. After air dried, the products were redissolved in 15 µL of formamide denaturing buffer. Sequencing the purified product, the nucleotide order was revealed, and these sequences were subsequently uploaded to the NCBI database. The GenBank database assigned the sequence an accession number. The biological sequence alignment editor BioEdit V 7.2.5 was used to trim the generated sequence.

RESULTS

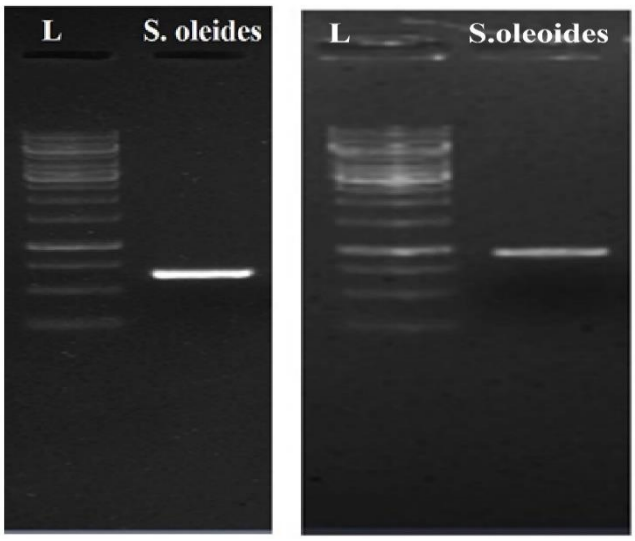
PCR analysis: The creation of DNA barcodes for species identification generally requires PCR-based amplification of conserved areas (*mat*K and *rbc*L). *rbc*L and *mat*K are both universal primers; we observed good results of PCR amplification *Salvadora oleoides.* Both primers resulted in efficient amplification (Fig. 3).



(a) (b)

Figure 3. PCR analysis of *Salvadora oleoides* (a) *rbc*L (b) *mat*K *L=1kb Ladder .

Gel Purification: The amplified PCR product was purified from gel using the FavorPrep^{TM} Gel purification mini kit (Fig. 4). The PCR purification technique removes short primers, dNTPs, enzymes, short-failed PCR products, and salts from PCR fragments of more than 100 bp quickly and effectively, usually in less than 10 minutes. After which, the purified product was checked on 1% agarose gel and, later on, sent for sequencing.



(a) (b)

Figure 4. Eluted products of *Salvadora oleoides* (a) *rbc*L, (b) *mat*K *L= 1kb Ladder .

Sequencing: Sequence homology of the amplified sequences was detected using Basic Local Alignment Tool (BLAST). For matK, the amplified product couldn't get sequenced due to the high heterozygosity, which might have resulted in polymerase slippage during Sanger sequencing method. The sequence length of rbcL was 526 nucleotides.

The sequenced sections revealed conserved genomic data that will be used to identify various plants that belong to these species in the future. After sequencing analysis, the amplified conserved barcodes indicated various degrees of biological similarity.

Bioinformatics studies: Different Bioinformatic analyses were performed on the obtained sequence. The sequence was submitted to GenBank after homology confirmation through the BLASTn tool that showed homology with *Salvadora persica*. The BLAST (basic local alignment search tool) program is used in bioinformatics to compare primary biological sequence data, the nucleotides of DNA or RNA sequences. The accession number assigned to *Salvadora oleoides* was OP046316 for *rbc*L.

The results showed that accurate species identification and discrimination could be achieved using these conserved DNA sequences as barcode primers.

DISCUSSION

The discovery of several genes and their intricate inheritance patterns in numerous plant species was made possible by the genome sequence analyses of numerous species and the mapping of complicated features linked with divergent phenotypes. DNA quality has a significant impact on PCR-based amplification. When metabolites are present in plants, they might occasionally affect the DNA quality during isolation, necessitating alternative DNA isolation techniques even for closely similar species (Khanuja *et al*., 1999). The fundamental method for determining a species' identity in plants is phylogenetic reconstruction and sequencing divergence from the reference sequence. Using DNA barcodes in native animals was a pioneering move toward developing DNA-based monitoring techniques for the adulteration of medicines in domestic and international commerce. However, many plant genomes are still missing sequencing information (Altschul *et al*., 1997). When identifying species of native land plants, a genus-based identification approach would be the best option. To find more conserved genomic areas among various plant species, re-sequencing more loci for target-based enhancements may be beneficial. The rbcL gene, which codes for a part of the crucial photosynthesis enzyme ribulose bisphosphate carboxylase, is found in almost all plant species.

In the current study, genes were amplified and sequenced in the conserved regions in the plant species *Salvadora oleoides*. Its DNA sequence has a region that varies greatly between

species, making it perfect for DNA barcoding. The MaturaseK gene (*MatK*) of chloroplast is highly conserved in plant systematics and is involved in Group II intron splicing (Notredame *et al*., 2000). The gene contains high substitution rates within the species and is emerging as a potential candidate to study of plant systematics and evolution. So, this gene was also used for the DNA barcoding of *Salvadora oleoides.* The samples were collected from different regions of Pakistan. From the young leaves of *Salvadora oleoides,* DNA extraction was done by the modified CTAB method.

DNA barcoding of endangered plant species *S. oleoides* has been obtained in this study using the two most reliable and efficient DNA barcodes *rbc*L and *mat*K. Due to the considerable sequence diversity in the primer binding sites, amplification and sequencing of the *mat*K barcoding region are challenging (Hollingsworth *et al*., 2011). Given that numerous studies have reported frequent PCR failures using these primers, the highly variable *mat*K area has low PCR amplification success than the more conserved *rbc*L gene (Kress and Erickson, 2007; Gonzalez *et al*., 2009). In this study, *mat*K showed 100% amplification result but the sample could not get sequenced. A similar result for the studies in which *mat*K cannot get sequenced were on *Durio graveolent* and *Durio zibetinus* (Cahyaningsih *et al*., 2021) and plants of the genus Acacia (Ismail *et al*., 2020).

In this research study, the species identification of *S. oleoides* has proven the strength of *rbc*L gene as the potential barcode. But, *mat*K showed less reproducibility. After the genes' sequences had been obtained, the signature DNA barcodes of this species were submitted to the NCBI. This study has proven the strength of *rbc*L as a suitable DNA barcode for the species identification of endangered species *Salvadora oleoides*.

***Conclusion*:** A unique *rbc*L sequence was found by amplifying and sequencing conserved genomic regions in engendered plant species of *Salvadora oleoides.* The study's conclusions might be used to develop DNA-based methods for classifying medicinal plant species and detecting adulteration.

***Funding*: N/A**

***Conflict of interest*:** The submitted research article is based on the original work, not been previously published and is not being considered for publication elsewhere at this time. Additionally, there are no competing interests.

***Ethical statement*:** This article does not contain any studies with human participants or animal performed by any of the authors.

***Availability of data and material*:** We declare that the submitted manuscript is our work, which has not been published before and is not currently being considered for publication elsewhere.

***Acknowledgement*:** Thanks to Plant Disease Resistance Lab" of CABB for providing materials for the study.

***Code Availability*:** Not applicable

***Authors' contributions*:** All the listed authors conducted the experiments and completed the, manuscript.

***Consent to participate*:** All authors are participating in this research study.

***Consent for publication*:** All authors are giving the consent to publish this research article in PDC.

REFERENCE:

Altschul, S.F., T.L. Madden, A.A. Schäffer, J. Zhang and Z. Zhang. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25:3389-3402.

Cahyaningsih, Y. F., A. Sutanto, P. J. Santoso, M. A.Arsyad and S. H. Larekeng. 2021. Amplification rate of *mat*K and rbcL genes in three types of durian. In *IOP* Conference Series: Earth and Environmental Science 886:012011.

Chen, S.L., H. Yu, H.M. Luo, Q. Wu, C.F Li, and A. Steinmetz. 2016. Conservation and sustainable use of medicinal plants: problems, progress, and prospects. Chinese medicine 1:1-10.

Dong, W., T. Cheng, C. Li, C. Xu, P. Long, C. Chen and S. Zhou. 2014. Discriminating plants using the DNA barcode *rbc* L : an appraisal based on a large data set. Molecular Ecology Resources. 14:336-343.

Ghorbani, A., Y. Saeedi and H.J. De Boer. 2017. Unidentifiable by morphology: DNA barcoding of plant material in local markets in Iran. PloS one 12:0175722.

Gonzalez, M.A., C. Baraloto, J. Engel, S.A. Mori and P. Pétronelli. 2009. Identification of Amazonian trees with DNA barcodes. PLoS ONE 4:7483.

Hebert, P. D., A. Cywinska, S. L.Ball, and J. R. DeWaard. 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences 270:313-321.

Hollingsworth, P. M., S. W. Graham and D. P. Little. 2011. Choosing and using a plant DNA barcode. PloS one 6:19254.

Ismail, M., A. Ahmad, M. Nadeem, M. A. Javed, S. H Khan, I. Khawaish, and S. Qamer. 2020. Development of DNA barcodes for selected Acacia species by using *rbc*L and matK DNA markers. Saudi Journal of Biological Sciences 27:3735-3742.

Khan, A.U. 1994. History of decline and present status of natural tropical thorn forest in Punjab. Biological Conservation 67:205-210

Khan, A.W., A.U. Khan, S.M.M. Shah, A. Ullah, M. Faheem and M. Saleem. 2019. An updated list of neuromedicinal plants of Pakistan, their uses, and phytochemistry. Evidence-Based Complementary and Alternative Medicine.

Khan, M. F., Hussain, and S. Musharaf. 2012. Conservation Status of Tree species in Tehsil Takht-e-Nasrati, Karak Pakistan. Journal of Biodiversity and Environmental Science 2:48-56

Khanuja S.P., A.K. Shasany, M.P. Darokar and S. Kumar. 1999. Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. Plant Molecular Biology Reporter 171:74-74.

Kress, W.J., and D.L. Erickson. 2007. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* apacer region. PLoS ONE 2:508.

Li, M., H. Cao, P.H But and P.C. Shaw. 2011. Identification of herbal medicinal materials using DNA barcodes. Journal of Systematics and Evolution 49:71-283.

Maloukh, L., A. Kumarappan, M. Jarrar, J Salehi, H. El-Wakil, and T.V. Rajya Lakshmi. 2017. Discriminatory power of rbcL barcode locus for authentication of some of United Arab Emirates (UAE) native plants. Biotechnology 7:1-7.

Notredame, C., D. G. Higgins and J. Heringa. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. Journal of molecular biology 302:205-217.